A Second Order Defect Correction Scheme for Unsteady Problems

Regis Martin, Hervé Guillard

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Second Order Defect Correction Scheme for Unsteady Problems

Régis Martin, Hervé Guillard

N° 2447

Novembre 1994

PROGRAMME 6

Calcul scientifique,

modélisation

et logiciel numérique

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

A Second Order Defect Correction Scheme for Unsteady Problems

Régis Martin, Hervé Guillard

Programme 6 — Calcul scientifique, modélisation et logiciel numérique
Projet SINUS

Rapport de recherche n ° 2447 — Novembre 1994 — 41 pages

Abstract: Using the defect correction method (DeC), we propose an implicit scheme that is second order accurate both in time and space,but that uses only first order jacobian. In a first time, we theoretically analyse the truncation error of the scheme and perform a linear stability analysis of it. Then some numerical experiments over simple test-cases are presented. Finally, the capability and accuracy of this new scheme is outlined by the analisys of a complex unsteady flow in a 2-D model of a piston engine.

Key-words: Computational Fluid dynamics, Implicit schemes, Defect Correction method

(Résumé : tsvp)

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Un schéma d'ordre 2 avec la méthode du défaut corrigé pour des problèmes instationnaires

Résumé : En utilisant la méthode du défaut corrigé (DeC), nous proposons un schéma implicite précis à l'ordre 2 en temps et en espace, mais n'utilisant que des jacobiens d'ordre 1. Dans un premier temps, nous analysons l'erreur de troncature du schéma puis démontrons sa stabilité linéaire inconditionnelle. On présente plusieurs expériences numériques sur des cas-tests simples. Un dernier calcul représentant un écoulement complexe instationnaire dans un modèle bi-dimensionnel de moteur à piston illustre les possibilités et la précision du schéma.

Mots-clé : Mécanique des fluides numérique, Schémas implicites, Méthode du Défaut corrigé

Contents

List of Figures

1 Introduction

Nowadays the most commonly employed discrete methods to solve steady hyperbolic problems use second-order discretization schemes. This is achieved either by central differencing [2, 7] or instead by second order upwind schemes (e.g. [4, 16]). The second alternative combined with TVD or MUSCL non linear dissipation has gained a large popularity among the CFD community. Most of the algorithms employed to implement these second order upwind schemes use a pseudo-time-marching procedure : Given a second order upwind discretization $\Phi^2(W_h)$ of the steady partial differential equation :

$$div(F(W)) = 0 \tag{1}$$

The approximate solution is sought as the limit when $j \to +\infty$ of the solution of the following system where τ is a pseudo-time step :

$$\frac{W_h^{j+1} - W_h^j}{\tau} + \partial_{W_h}\Phi^2(W_h^j)(W_h^{j+1} - W_h^j) = -\Phi^2(W_h^j) \tag{2}$$

In the limit of the large time step $\tau \to +\infty$, it is clear that this last approach reduces to Newton's method. However, the use of the system (2) is often unpractical. First, these approaches require the computation of the jacobian matrix $\partial_{W_h}\Phi^2$ which is rather difficult to express in closed form if possible. Actually, even an approximate linearization of this function may be quite involved computationally. Second, the linear systems (2) needs to be solved. This is computationally difficult and expensive.

These reasons motivated the use of alternate strategies into which the matrices appearing in the left hand side of (2) are replaced by computing a crude approximation of $\partial_{W_h}\Phi^2$. Of particular interest is the case where this approximation is choosen as a first-order upwind approximation of the continuous problem $div(F(W)) = 0$ leading to an iterative method of the form :

$$\frac{W_h^{j+1} - W_h^j}{\tau} + \partial_{W_h}\Phi^1(W_h^j)(W_h^{j+1} - W_h^j) = -\Phi^2(W_h^j) \tag{3}$$

It is clear that (3) provides at convergence a second -order accurate approximation of the solution of the steady system $div(F(W)) = 0$. The only

modification with respect to (2) concerns the sequence of successive iterates W_h^j. Computationally the loss of the quadratic convergence of Newton's method is counterbalanced by the reduction of the bandwidth of the matrices and the ease in solving the linear systems. Actually and at least for model problems, it can be shown that $\partial_{W_h}\Phi^1$ is very often diagonally dominant. Therefore its inversion can be performed by simple iterative methods as Gauss-Seidel or Jacobi relaxations. In the limit $\tau \to +\infty$, the convergence properties of scheme (3) have been studied in [3] where it was shown that for an appropriate choice of the discretizations $\Phi^1(W_h)$ and $\Phi^2(W_h)$, scheme (3) exhibits a good convergence rate.

We consider now the case where one is interested by the unsteady version of (1) :

$$\frac{\partial}{\partial t}W + div(F(W)) = 0 \tag{4}$$

In this case (3) is no more a second-order approximation of the unsteady system. Although consistent, this scheme is only formally first-order accurate and one may question its capability to compute really unsteady phenomena with large time steps. Actually, computational evidences have shown that the solution of (3) is largely better than plain first-order schemes and that this scheme can capture complicated unsteady phenomena that true first order schemes failed to compute [9, 10]. However there may be some cases where a true second-order accuracy is desirable and one may want to look for a more accurate approximation than (3) is.

The purpose of this paper is to show theoretically and numerically, by using the Defect-Correction (DeC) theory, that (3) can still be the basis of a true second-order accurate approximation of (4). The resulting scheme that avoids the tedious and often unpractical computation of the jacobian matrix $\partial_{W_h}\Phi^2$ is computationally attractive with good stability properties. The summary of this paper is as follows : In Section 2, we make precise the definition of the second-order scheme we propose and study some of its properties. In particular, we establish that the proposed method is second-order accurate in time and space and perform a linear stability analysis of this scheme. Section 4 is

devoted to numerical tests on some model problems and we end this paper by a description of a bi-dimensional problem where the approximation computed by (3) is compared with the improved one computed by the scheme we propose.

2 A Second Order DeC scheme for unsteady problems

We consider an evolution equation that we write in an abstract way :

$$\text{Find } W \in \mathcal{C}^0([0,T];H) \text{ such that :}$$

$$\begin{array}{rcl} \dfrac{\partial}{\partial t}W + F(W) & = & 0 \\ W(t=0) & = & W_0 \end{array} \tag{5}$$

where H is an Hilbert space, $W_0 \in H$ and F is a partial differential operator whose domain $D(H)$ is included in H. A semi-discrete scheme is obtained by defining a finite-dimensional space V_h included in H and an operator Φ_h from V_h to V_h, resulting in the ordinary differential system :

$$\text{Find } W_h \in \mathcal{C}^0([0,T];V_h) \text{ such that :}$$

$$\begin{array}{rcl} \dfrac{d}{dt}W_h + \Phi_h(W_h) & = & 0 \\ W_h(t=0) & = & W_{0,h} \end{array} \tag{6}$$

As usual, h represents a spatial discretisation step such that dim $V_h \to \infty$ when $h \to 0$. We also suppose that there exists a stable restriction R_h from H into V_h such that :

$$\lim_{h\to 0} ||R_h W - W||_H = 0 \quad \forall W \in H \quad \text{and} \quad ||R_h|| = \sup_{||W||=1} ||R_h W||_H \leq C$$

The operator Φ_h^p will be said to be accurate of order p if the truncation error defined by $T_h(W) = ||R_h F(W) - \Phi_h(R_h W)||_H$ is such that :

$$T_h(W) < C(W)h^p \quad \forall W \in \mathcal{D} \tag{7}$$

where $\mathcal{D}$ is a dense subset of H (Usually, $\mathcal{D}$ is a set of sufficiently differentiable functions). A completely discrete scheme is obtained from (6) by approximating the time derivative. In a similar way, we will said that a time discretization formula :

$$\frac{d}{dt}W \sim \frac{aW(t+\tau) - V^*(W(t), W(t-\tau), \ldots, W(t-j\tau))}{\tau} \tag{8}$$

where $V^*(.,\ldots,.)$ is a j-linear function, is p-order accurate if the truncation error is a term of order p :

$$||\frac{d}{dt}W - \frac{aW(t+\tau) - V^*}{\tau}|| < C(W)\tau^p \tag{9}$$

In the sequel, we assume that there is a relationship between the time and space steps of the type $\tau = rh$ with $r \geq 1$. The completely discrete approximation we will start with, writes :

$$\frac{aW_h^{n+1} - V^*(W_h^n, \ldots, W_h^{n-j})}{\tau} + \Phi_h^p(W_h^{n+1}) = 0 \tag{10}$$

We denote by $G_h(\tau)$, the operator from V_h into V_h that associate to $V^*(W_h^n, \ldots, W_h^{n-j})$, W_h^{n+1} solution of the non linear algebraic equation (10). The local truncation error associated to scheme (10) is defined to be :

$$T_h = \frac{R_h W(t+\tau) - G_h(\tau) R_h V^*(W(t), \ldots, W(t-j\tau))}{\tau} \tag{11}$$

and it is easy to see from (7), (9) that T_h is a term of order p, provided the operator $(a + \tau\phi_h^p)^{-1}$ is uniformely bounded.
Nevertheless, solving (10) is a formidable task : The jacobian $\partial_W \phi_h^p$ (if it exists !) can be very difficult to compute in analytical form and solving the linear systems $(a + \tau\partial_W \phi_h^p)X = b$ be computationally so intensive that a direct application of Newton method would be prohibitively expensive. Thus instead, given an m^{th} order ($m \leq p$) approximation of the operator F, such that $\partial_W \phi_h^m$ is easy to invert, we consider solving (10) by the modified Newton method (Defect Correction scheme) :

$$\begin{aligned} \mathcal{W}_h^0 &= \mathcal{W}^0(W_h^n, \ldots, W_h^{n-j}) \\ \mathcal{W}_h^{s+1} &= \mathcal{W}_h^s - \tau\,(a + \tau\partial_w \phi_h^m)^{-1} \left(\frac{a\mathcal{W}_h^s - V^*(W_h^n, \ldots, W_h^{n-j})}{\tau} + \phi_h^p \mathcal{W}_h^s \right) \end{aligned} \tag{12}$$

each step of (12) defines an operator $G_h^s(\tau)$ that associate to V^*, the vector $\mathcal{W}_h^s$. Denoting by T_h^s, the associated truncation error, it is easy to see from (12) that the truncation error T_h^s obeys the recursion formula :

$$T_h^{s+1} = A(T_h^s - T_h) + T_h \tag{13}$$

where $A = I - (a + \tau\partial_W \phi_h^m)^{-1}(a + \tau\phi_h^p)$.

Solving for T_h^s gives :

$$T_h^s = A^s(T_h^0 - T_h) + T_h \tag{14}$$

where T_h^0 is the truncation error associated with the initialisation of the Newton iteration (12) :

$$T_h^0 = \frac{R_h W(t+\tau) - R_h \mathcal{W}^0(W(t), \ldots, W(t-j\tau))}{\tau} \tag{15}$$

If $\rho(A) < 1$, (14) shows that the truncation error $T_h^s \to T_h$ which is no surprise. However, it also shows that there is no need to converge the DeC iterations (12) to reach p-order accuracy, it suffices to perform a finite number of iterations s_p such that the first term of (14) is of order p, i.e $s_p = Ln(T_h/T_h^0)/Ln\rho$. This principle sometimes called the finite termination property of DeC schemes make them particularly attractive. The same principle is also applied in the so-called Full Multi-grid techniques (FMG see [6]). We now proceed to evaluate s_p. We first remark :

Lemma 1 : $\exists K_\Delta$ independent of h and V such that for any $V \in \mathcal{D}$ (i.e sufficiently differentiable) we have (at least) :

$$||(a + \tau\partial_w \phi_h^m)^{-1}(\phi_h^p - \phi_h^m) R_h V||_H \leq K_\Delta h^m ||D^{m+1} V||_H \tag{16}$$

Proof : This follows from the triangle inequality, that ϕ_h^m and ϕ_h^p are order m and p ($m < p$) approximations of the same operator and that $(a+\tau\partial_w\phi_h^m)^{-1}$ form a family of bounded operator for $h \to 0$ (we assume of course that the m-order implicit linearized scheme is stable). □

Bounding now, $||T_h^s||$ by $||A^s T_h^0||+2||T_h||$, noting that $A = \tau(a+\tau\partial_w\phi_h^m)^{-1}(\phi_h^m - \phi_h^p)$ and using lemma 1, we obtain :

$$||A^s T_h^0|| \leq K_\Delta^s r^{s-1} h^{(m+1)s-1} ||D^{(m+1)s}(W(t+\tau) - \mathcal{W}^0)|| \tag{17}$$

Then suppose that $\mathcal{W}^0$ is an approximation of order l of $W(t+\tau)$ i.e :

$$W(t+\tau) = \mathcal{W}^0 + \alpha.\tau^l \partial_t^l W$$

(17) implies :

$$||A^s T_h^0|| \leq K_\Delta^s \alpha r^{s+l-1} h^{(m+1)s+(l-1)} ||D^{(m+1)s}\partial_t^l W|| \tag{18}$$

and thus this establishes :

Proposition 1 : The truncation error T_h^s is at least of order p as soon as $s_p \geq (p-l+1)/(m+1)$.

Remark : The truncation error T_h (11) of the fully implicit non-linear scheme is bounded by an expression of the form :

$$||T_h|| \leq K h^p \left(||D^{(p+1)}W|| + r^p ||\partial_t^p W|| \right) \tag{19}$$

we observe that if $s_p \geq (p-l+1)/(m+1)$, not only the formal order of accuracy of the two component of the global truncation order (the exponent of h) are equal but also that the order of the derivatives involved in (18) and (19) are similar.

One of the most interesting application of the above theory concerns the case where ϕ_h^m is a first-order upwind approximation of a hyperbolic first-order operator F. Specializing the discussion to this case, we have :

Corollary 1 : Let ϕ_h^2 be a second-order approximation of F, then :
if $\mathcal{W}^0 \equiv 0$, $s_p \geq 2$ is required to obtain a second-order accurate scheme.
if $\mathcal{W}^0 \equiv W^n$, $s_p \geq 1$ suffices to obtain a second-order accurate scheme.

Thus using $s_p = 2$ with $\mathcal{W}^0 \equiv W^n$ does not improve the formal order of accuracy of the scheme, neither does the *a priori* more accurate initialization

$\mathcal{W}^0 = 2W^n - W^{n-1}$.

For completness, if one is interested in third-order accurate schemes, we can similarly prove :

Corollary 2 : Let ϕ_h^3 be a third-order approximation of F, then :
if $\mathcal{W}^0 \equiv W^n$, $s_p \geq 2$ is required to obtain a third-order accurate scheme.
but if $\mathcal{W}^0$ is a second-order approximation of $W(t+\tau)$, (for instance $\mathcal{W}^0 = 2W^n - W^{n-1}$) $s_p \geq 1$ would suffice to obtain a third-order accurate scheme.

Remark 2 : The above two corollaries shows that second and third order accurate implicit schemes can be obtained at the price of inverting *only once* the jacobian matrix arising from a *first order* discretization.

We now proceed to test numerically the above theory by applying it to the scalar advection equation :

$$\frac{\partial}{\partial t}u + \frac{\partial}{\partial x}u = 0$$

with periodic boundary conditions. To construct a second-order accurate implicit scheme, we will use the following spatial approximation of the first-order derivative :

$$\begin{aligned} \phi^1(jh) &= \frac{u(jh) - u(jh-h)}{h} \\ \phi_c^2(jh) &= \frac{u(jh+h) - u(jh-h)}{2h} \quad \phi_u^2(jh) = \frac{3u(jh) - 4u(jh) + u(jh-2h)}{2h} \end{aligned} \tag{20}$$

and $\phi^2 = 1/2(\phi_c^2 + \phi_u^2)$.

The time discretization scheme uses the second-order backward differentiation formula:

$$\frac{\partial}{\partial t}u \sim \frac{3u^{n+1} - 4u^n + u^{n-1}}{2\tau}$$

In the computation reported below the CFL number equals 10. Figure 1 displays the slopes of the error curves $||u_{exact} - u_h||_{L^2}$ at time $t = 1.5$ as the grid is raffined. We use the following notations : NcV denotes the second-order

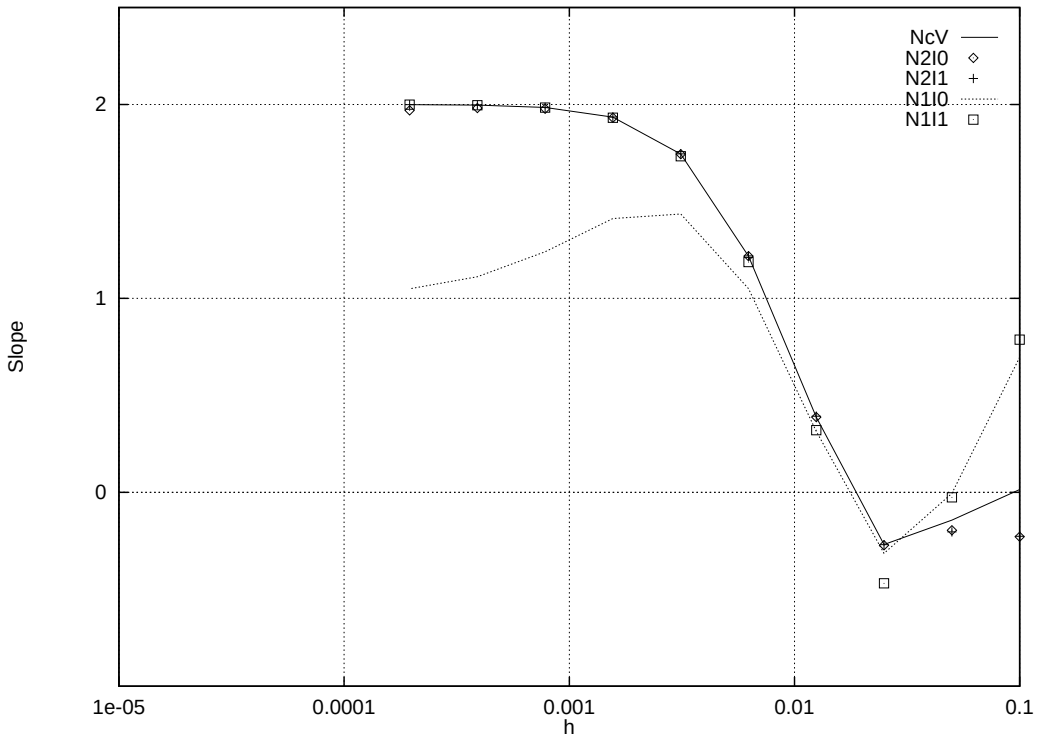


Figure 1: Slope of the error curve $Cfl = 10$

scheme obtained after full convergence of (12), N1I1 (one Newton step, initialisation of first order) the scheme (12) after one step with $\mathcal{W}^0 = W^n$, N2I1 the scheme (12) after two steps with $\mathcal{W}^0 = W^n$ while N1I0 and N2I0 will denote the schemes obtained after respectively one and two Newton steps with the 0-order discretization $\mathcal{W}^0 = 0$.

The slope of the error curve corresponding to NcV tends to 2 as $h \to 0$. As predicted by the above theory, it can be checked that this is also the case for the N1I1 and N2I1 schemes. The slope of the error curve $\to 1$ for the scheme N1I0 while in perfect agreement with Corollary 1, second order accuracy is recovered by a second Newton step (N2I0 scheme).

3 Stability Analysis

In this section, we concentrate on the second order Dec schemes N1I1 and N2I1 and perform a stability analysis on the linear model equation :

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \quad \forall x \in R \quad \forall t > 0 \tag{21}$$

Given a regular mesh $x_j = j.h$, we define $\mathcal{U}_h^n = \{u(x_j, t^n)\}_{j \in Z}$ and initialize the DeC iterations (12) by $\mathcal{U}^0 = \eta_1.U^n + \eta_2.U^{n-1}$. With these notations, (12) reduces to :

$$H_1(\mathcal{U}^{l+1} - \mathcal{U}^l) = -H2.\mathcal{U}^l + V^* \tag{22}$$

where $H_1 = [3/2\ I + \tau\phi_h^1]$, $H_2 = [3/2\ I + \tau\phi_h^2]$ and $V^* = -2.U^n + 1/2\ U^{n-1}$. We can then express $\mathcal{U}^l$ by :

$$\left\{ \begin{array}{lcl} \mathcal{U}^l & = & A^l.\mathcal{U}^0 + (I - A^l)H_2^{-1}V^* \\ A & = & I - H_1^{-1}H_2 \end{array} \right. \tag{23}$$

To analyse the stability of (22) we rewrite it in the following way :

$$\begin{pmatrix} \mathcal{U}^l \\ U_h^n \end{pmatrix} = G_h^l \begin{pmatrix} U_h^n \\ U_h^{n-1} \end{pmatrix}, \quad G_h^l = \begin{pmatrix} G_1^l & G_2^l \\ Id & 0 \end{pmatrix}$$

where

$$\begin{array}{l} G_1^l = \eta_1.A^l - 2.(A^l - Id).H_2^{-1} \\ G_2^l = \eta_2.A^l + 1/2\ (A^l - Id).H_2^{-1} \end{array}$$

The spectral radius of G_h^l is computed with the help of Fourier analysis. Introducing a Fourier mode $E_k = (..., e^{ik(j-1)h}, e^{ikjh}, e^{ik(j+1)h}, ...)$ into (23) :($U_h^n = \hat{U}_k^n.E_k$, $U_h^{n+1} = \hat{U}_k^{n+1}.E_k$, $U_h^{n-1} = \hat{U}_k^{n-1}.E_k$) and using that the eigenvalues of ϕ_h^1 and ϕ_h^2 are respectively $d^1(k)$ and $d^2(k)$:

$$\begin{array}{rcl} d^1(k) = & & \left(1 - e^{-ikh}\right)/h \\ d^2(k) = & & \left(e^{-i2kh} - 5e^{-ikh} + 3 + e^{ikh}\right)/(4h) \end{array} \tag{24}$$

we obtain after some algebra, that the coefficients of the amplification matrix G_k^l defined by :

$$\begin{pmatrix} \hat{\mathcal{U}}_k^l \\ \hat{U}_k^n \end{pmatrix} = G_k^l \begin{pmatrix} \hat{U}_k^n \\ \hat{U}_k^{n-1} \end{pmatrix} \quad , \quad G_k^l = \begin{pmatrix} g_1^l(k) & g_2^l(k) \\ 1 & 0 \end{pmatrix} \tag{25}$$

are given by :

$$\begin{aligned} g_1^l(k) &= \eta_1 \lambda_A^l(k) - 2\frac{(\lambda_A^l(k)-1)}{\lambda_{H2}(k)} \\ g_2^l(k) &= \eta_2 \lambda_A^l(k) + \tfrac{1}{2}\frac{(\lambda_A^l(k)-1)}{\lambda_{H2}(k)} \end{aligned} \tag{26}$$

with :

$$\begin{aligned} \lambda_A(k) &= \frac{\lambda_{H1}(k) - \lambda_{H2}(k)}{\lambda_{H1}(k)} \\ \lambda_{H1}(k) &= 3/2 + \mu(h.d_1(k)) \\ \lambda_{H2}(k) &= 3/2 + \mu(h.d_2(k)) \end{aligned}$$

where $\mu = \tau/h$ represents the CFL number. The stability of the scheme is insured if $||G_k^l||$ is uniformely bounded. This can be checked by applying the result cited in Richtmyer and Morton book ([12], p. 86) requiring that all the eigenvalues of G_k^l strictly lie inside the unit circle, excepted one maximum which norm can equal one. For the schemes N1I1 and N2I1, this is easily checked by using Miller s theorem (see e.g. [11]) or instead by computing directly these eigenvalues: The "spurious root " (the eigenvalue approximating 0) is bounded in norm by 1/3, and the "physical root" (the one approximating $e^{-i\tau}$) is always inside the unit cicle, excepted for $k = 0$ where it is equal to one.

Figures 2 and 3 show the norm of these eigenvalues, computed for a CFL number equals to 50 as function of the parameter $\theta = kh$. Figure 2 displays the "physical" eigenvalue while figure 3 shows the "spurious" one.

We note that for low frequency, the behavior of the two schemes is very similar. For $k = 0$ (i.e. $\theta = 0$) the physical eigenvalues are equal to 1 as expected. For the spurious roots (fig 3) the absolute value of the eigenvalues

corresponding to the null mode is equal to $1/3$ as can be checked directly from (26). For the whole range of frequencies, the norms of the spurious root of the two schemes are very similar. The largest differences between the two schemes N1I1 and N2I1 appear for $\theta \sim \pi/2$ and for the physical eigenvalues. It can be checked from figure 2 that for these frequencies, the physical amplification factor of scheme N1I1 is around $1/2$. This behavior persits as the CFL number increases and actually, it can be shown that for CFL $\rightarrow +\infty$, the amplification factor $\rightarrow 1/2$ as previously shown in ([3]). This behavior reflects the "inconsistency" between the first order discretisation ϕ_h^1 and the second-order one ϕ_h^2. Indeed, it is easily shown that the amplification factor of the plain second-order scheme tends to 0 as CFL $\longrightarrow \infty$. When one uses scheme N2I1, that is intermediate between N1I1 and plain second-order scheme, it can be seen from figure 2 that the amplification factor is less than $1/4$ in the hight frequency range.
This analysis leads us to the conclusion that although the two schemes N1I1 and N2I1 are both linearly unconditionally stable, scheme N1I1 presents a "pathological" behavior in the medium frequency range and that scheme N2I1 possesses better stability properties. In the sequel, we will numerically investigate the differences between these two second-order schemes.

Remark: The second order DeC scheme (12) N1I2 obtained from (12) by performing one Newton iteration and initializing with $\mathcal{W} = 2W^n - W^{n-1}$ is not stable, due to the spurious root whose norm grows larger than one. Thus, not only the initialization $\mathcal{W} = 2W^n - W^{n-1}$ does not improve the formal order of accuracy of the scheme, but this initialisation has also a negative effect on the stability.

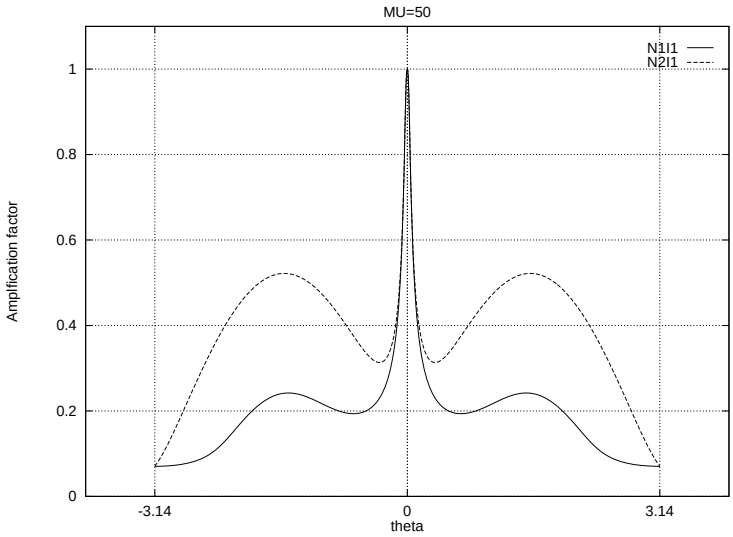


Figure 2: Physical eigenvalues with $\mathcal{U}^0 = U_h^n$ and $Cfl = 50$

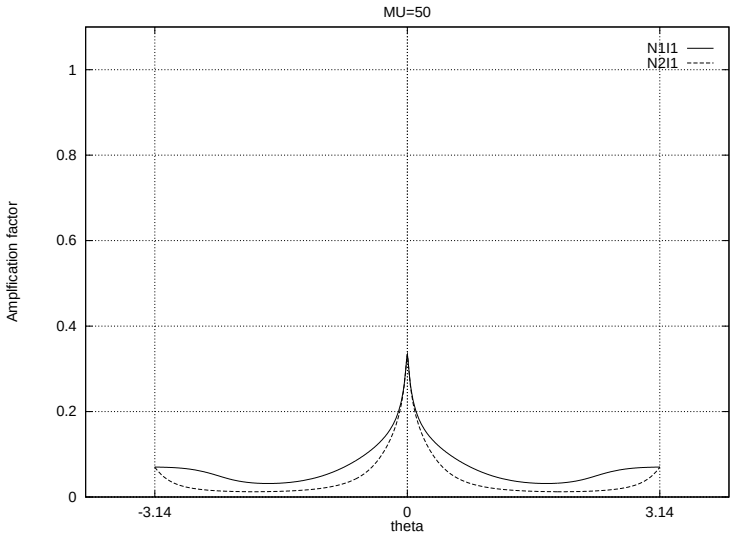


Figure 3: Spurious eigenvalues with $\mathcal{U}^0 = U_h^n$ and $Cfl = 50$

4 Numerical Tests

In this section,we are interested in applying the implicit second-order defect correction scheme to Euler's equation of fluid dynamics, that we define below :

$$\begin{cases} \frac{\partial}{\partial t}W + \frac{\partial}{\partial x}F(W) + \frac{\partial}{\partial y}G(W) &= 0 \\ W(x,y,t=0) &= W_0(x,y) \qquad \forall (x,y) \in \Omega \\ W(x,y,t) &= W_\infty(x,y,t) \quad \forall (x,y) \in \partial\Omega, \forall t > 0 \end{cases} \tag{27}$$

$$\text{with : } W = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho e \end{pmatrix}, F(W) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (\rho e + p)u \end{pmatrix}, G(W) = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (\rho e + p)v \end{pmatrix}$$

where $\rho, (u,v), e$ are respectively the density,velocity and total specific energy of the fluid. p represents the pressure, and is related to others thermodynamical variables by the perfect gas law :

$$p = (\gamma - 1)(\rho e - \frac{1}{2}\rho(u^2 + v^2))$$

4.1 Finite Volume Method on unstructured meshes

The spatial discretisation of this set of equation uses a mixed finite element/finite volume method that we describe rapidly below. For further details, we refer to ([1],[5]). We assume that Ω is a polygonal domain and we triangulate Ω by a finite element triangular mesh $\mathcal{T}_h$. For each node of the triangulation, a control volume $\mathcal{C}_i^h$ is defined by means of the medians of the edges of the triangles and the barycenters of the triangle as shown in Fig(4).

Integrating (27) over $\mathcal{C}_i^h$ leads to :

$$\frac{\partial}{\partial t}\int_{C_i^h} W\,dv + \sum_{j \in V(i)} \Phi_{ij} + BT = 0$$

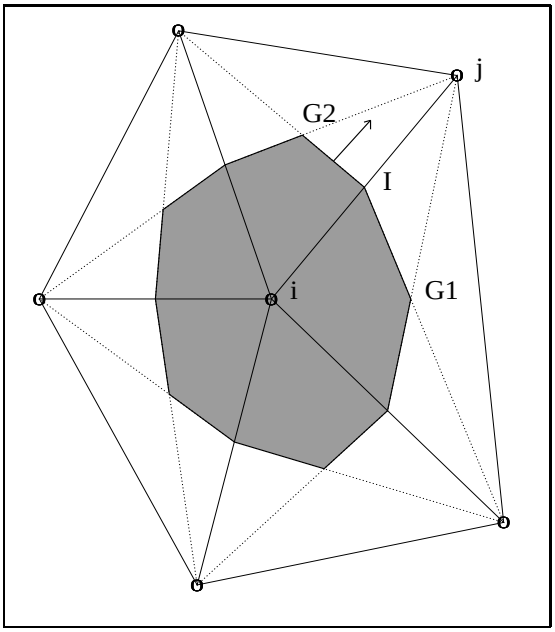


Figure 4: Control volume of $\mathcal{C}_i^i$

where $V(i)$ denotes the set of the neighboors of node i, and BT stands for the boundary fluxes. In the above edxpressions, Φ_{ij} represents the intercell fluxes defined by:

$$\Phi_{ij} = \int_{\partial C_{ij}^h} F.n_{ij}^x + G.n_{ij}^y ds$$

where ∂C_{ij}^h is the intersection of the cells $\mathcal{C}_i^h$ and $\mathcal{C}_j^h$, and $\vec{n_{ij}} = (n_{ij}^x, n_{ij}^y)$ is the outward unitary normal of ∂C_{ij}^h The approximation of the intercell flux Φ_{ij} uses a classical Roe approximate Riemann solver [13] that writes

$$\begin{array}{rcl} \Phi_{ij}^1 & = & \|\partial C_{ij}\| \mathcal{F}_{Roe}(W_i, W_j, \vec{n_{ij}}) \\ \mathcal{F}_{Roe}(W_i, W_j, \vec{n_{ij}}) & = & \frac{1}{2}\left[(F(W_i) + F(W_j)).n_{ij}^x + (G(W_i) + G(W_j)).n_{ij}^y\right] \\ & & +\frac{1}{2}|\overline{A}(W_i, Wj, \vec{n_{ij}})|(W_i - W_j) \end{array}$$

where $\overline{A}(W_i, Wj, \vec{n_{ij}})$ is the Roe matrix (see [13]). Φ_{ij}^1 is a first order accurate approximation of the Euler fluxes in the sense of (7). To obtain a second order accurate approximation, the classical MUSCL method by van Leer ([15]) is used.

$$\begin{array}{rcl} \Phi_{ij}^2 & = & \|\partial C_{ij}\| \mathcal{F}_{Roe}(W_{ij}^-, W_{ij}^+, \vec{n_{ij}}) \\ W_{ij}^- & = & W_i + \frac{1}{2}\nabla W_i.\vec{ij} \\ W_{ij}^+ & = & W_j - \frac{1}{2}\nabla W_j.\vec{ij} \end{array} \tag{28}$$

where ∇W_i, is the gradient of W at node i, evaluated by an average of the P_1 gradient of W on the triangles having i as node . The time discretisation uses the formula (8).

In summary in this case, the resulting implicit scheme N2I1 takes the form :

$$\begin{array}{rcl} \left[\frac{3}{2}\ Id + \tau\partial_W \Phi_h^1(W_h^n)\right](W - W_h^n) & = & -(\frac{3}{2}W_h^n + \tau\Phi_h^2(W_h^n)) \\ & & +(2\ W_h^n - \frac{1}{2}\ W_h^{n-1}) \\ \left[\frac{3}{2}\ Id + \tau\partial_W \Phi_h^1(W_h^n)\right](W_h^{n+1} - W) & = & -(\frac{3}{2}\ W + \tau\Phi_h^2(W)) \\ & & +(2\ W_h^n - \frac{1}{2}\ W_h^{n-1}) \end{array} \tag{29}$$

where $(\Phi_h^2)_i = \sum_{j \in V(i)} \Phi_{ij}^2$ and $\partial_W \Phi_h^1$ is an approximation of the jacobian of the first order fluxes defined by :

$$(\partial_W \Phi_h^1(U)\delta W)_{ij} = \frac{\left[A(U_i, \vec{n}_{ij}) + |\overline{A}(U_i, Uj, \vec{n}_{ij})|\right]}{2}.\delta W_i + \frac{\left[A(U_j, \vec{n}_{ij}) - |\overline{A}(U_i, Uj, \vec{n}_{ij})|\right]}{2}.\delta W_j$$

where $A(U_j, \vec{n}_{ij}) = \partial_W F(U_j) n_{ij}^x + \partial_W G(U_j) n_{ij}^y$ while the implicit scheme N1I1 results from the first step of (29).

4.2 Density wave convection

We consider a homogenous flow ($\rho = 1.4 g.cm^{-3}$, $P = 10^4 atm$) of constant speed ($10 cm.s^{-1}$) in a 2cm length canal. An initial perturbation of the density field ($\rho = 1.6 g.cm^{-3}$) is located between $x = 0.2$ and $x = 0.3$. The solution of this problem consists in the convection at the constant velocity $\vec{V} = 10 cm.s^{-1}$ of this discontinuity. We also note that in this case, Euler's equations reduces to the simple linear convection equation (21). We now compare the Defect Correction schemes N2I1 and N1I1.To discriminate between the effect of time and space accuracy, we also consider the first order implicit scheme (3). All the computations have been done by using the van Albada-van Leer limiters [14]. Figure (5) shows the results obtained with a CFL number of 10 (note that this correspond to a convective CFL, $u\tau/h$ of 1). It can be seen that the solutions obtained with the two schemes N2I1 and N1I1 are in good agreement with the exact solution and largely better than the solutions obtained with the first order scheme. The second defect Correction step of (29) does not seem to have a strong influence on the solution, except that we note that the oscillation at the rear of the discontinuity that is rather larger with N1I1 is smoothed out by the second order step of scheme N2I1. We have checked that further Newton iterations almost do not improve the results and that this oscillation is still present after convergence. Figure (6,7) show the results at a CFL number of 50 and 100 (ie a convective number of 5 and 10). The previous conclusions remains valid for these large time steps: the solutions obtained with the first order scheme is much more damped than the solutions obtained with N2I1

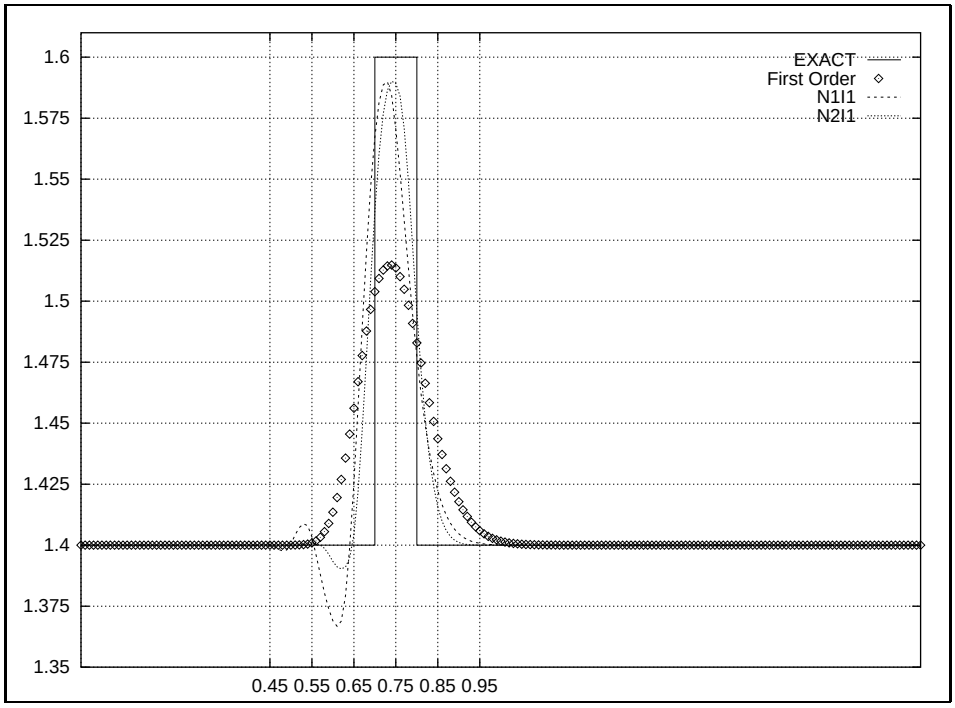


Figure 5: Density wave convection CFL=10

and N1I1, and the second order step contribute to a modest improvment of the solution, mainly in reducing the amplitude of the oscillation present at the rear of the density wave. This finding is in good qualitative agreement with the results obtained from the previous linear stability analysis indicating that scheme N2I1 is more dissipative that scheme N1I1 in the high frequency range.

4.3 Isentropic Compression

We next consider a problem well suited to the use of implicit schemes. This is a 1-D isentropic compression-expansion obtained by compressing a gas contained in a closed tube, closed at an end by a moving piston. In this experiment the

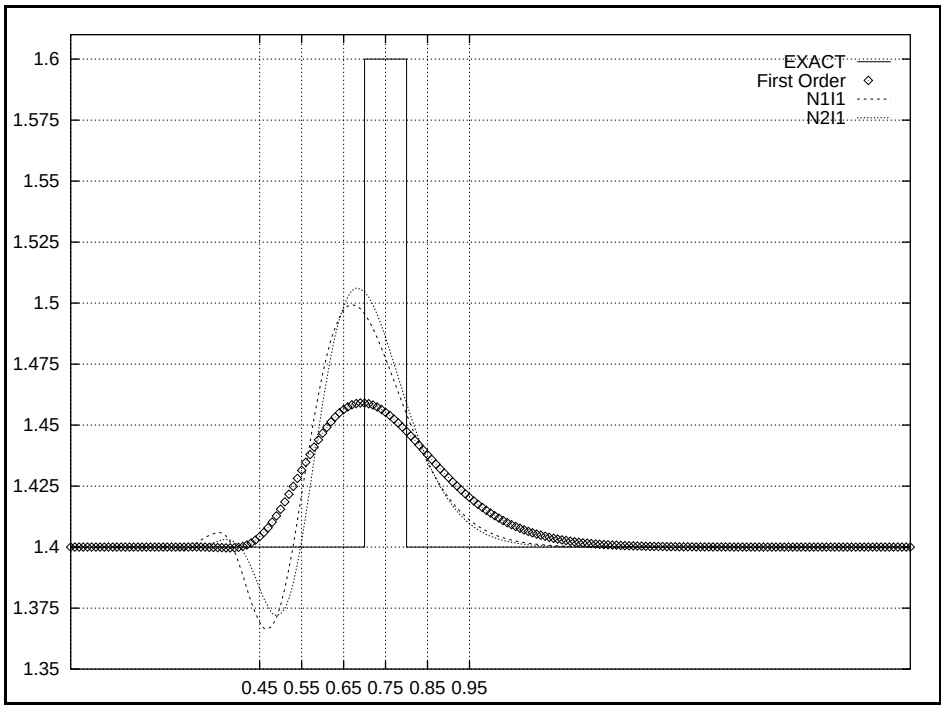


Figure 6: Density wave convection CFL=50

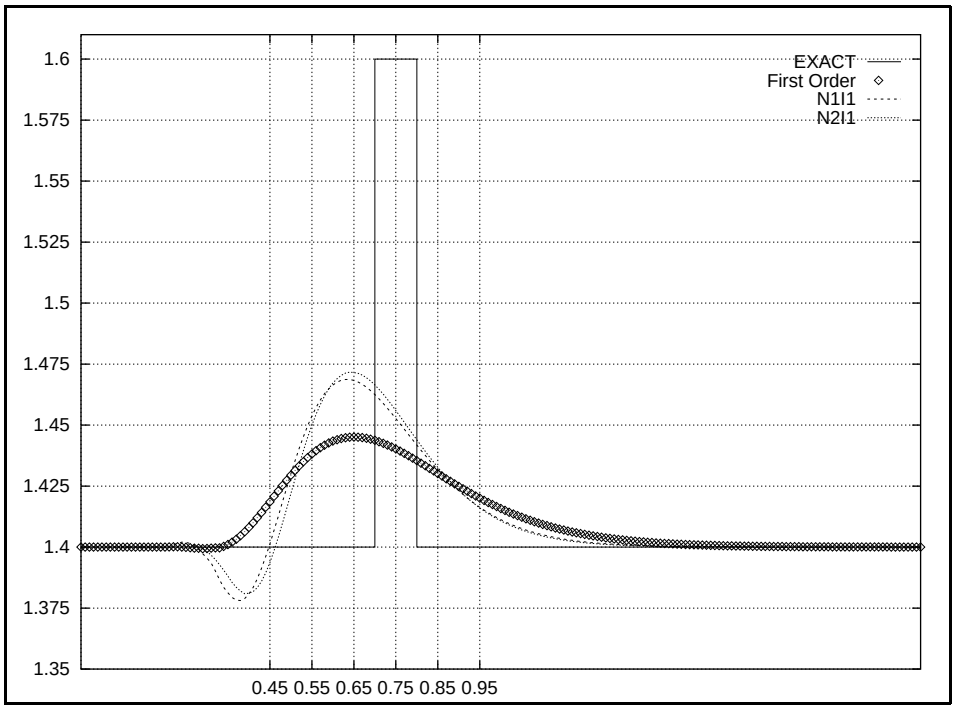


Figure 7: Density wave convection CFL=100

length of the tube is 9.032 cm and the gas is initially at rest with a density and pressure of respectively $1.19 g.cm^{-3}$ and $10^6 Pa$. At $t = 0$, the piston is moved in a smooth manner according to the law :

$$Xp(t) = (\frac{\alpha}{2} + \beta) - \frac{1}{2}\alpha cos(\theta(t)) - \sqrt{\beta^2 - \frac{\alpha^2 sin(\theta(t))^2}{4}} \tag{30}$$

where α, β are constant coefficients. This law mimics the movement of a piston moved by a crankshaft system and in this expression θ is the crank angle (CA) defined by :

$$\theta(t) = \theta_0 + 2\pi * (rpm/60)$$

where rpm is the number of rotation per minute. The values of the coefficient were here choosen to be :

$$\alpha = 8.9cm, \beta = 15.5cm, rpm = 2000tr/min$$

For these values, a typical Mach number of the solution will be around 3.10^{-2}. The use of implicit scheme to compute the solution of this type of problem is thus mandatory. By using the Low Mach number approximation, an analytical solution of this problem can be found (see e.g [10]) that shows that the velocity is linear, the pressure is a quadratic function and that density is almost constant. To take into account the displacement of the piston, the computation is performed on a moving mesh. This imposes some modifications in the expression of the fluxes. We refer to [8],[10] for a detailed description of the algorithm on a moving mesh. We show Fig (9,8) respectively, the reduced density $\tilde{\rho}$ and the reduced pressure $\tilde{p}$ defined by :

$$\tilde{\rho} = \frac{\rho(x,t) - \rho(0,t)}{Max_x \rho(x,t)}$$

$$\tilde{p} = \frac{p(x,t) - p(0,t)}{Max_x p(x,t)}$$

The computations have been done on a 21x3 node mesh, with a constant time step corresponding to a variation of 0.5 degres of the parameter θ in formula (30).The CFL number goes from 5 (at the beginning of the computation) to

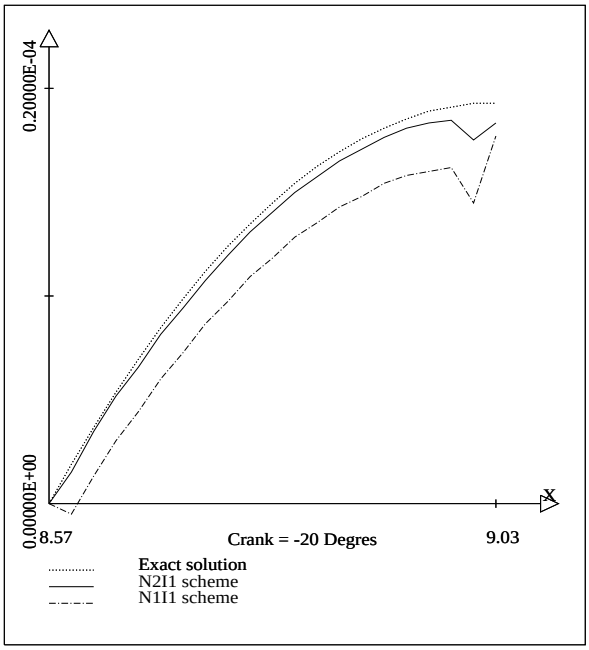


Figure 8: Isentropic compression: pressure

120 (at the end). One notice the large improvment produced by the second-order DeC N2I1: The pressure is closer to the analytical solution with N2I1 than with N1I1. This improvment is emphasized by considering the density curves where the solutions obtained with N2I1 scheme is almost the same than the exact one. We also note on these density curves that N2I1 scheme reduces dramatically the oscillations of the density present near the piston.

4.4 Sod Shock Tube

Our third numerical test is composed of the classical Sod shock tube problem. We emphasize that this strongly unsteady problem is not well suited for implicit schemes at large CFL number, and that the overall performances of implicit schemes on this problem are poor. The computations have been performed in a square of 1cm length on a 101x3 node mesh with a CFL number of 5. The van

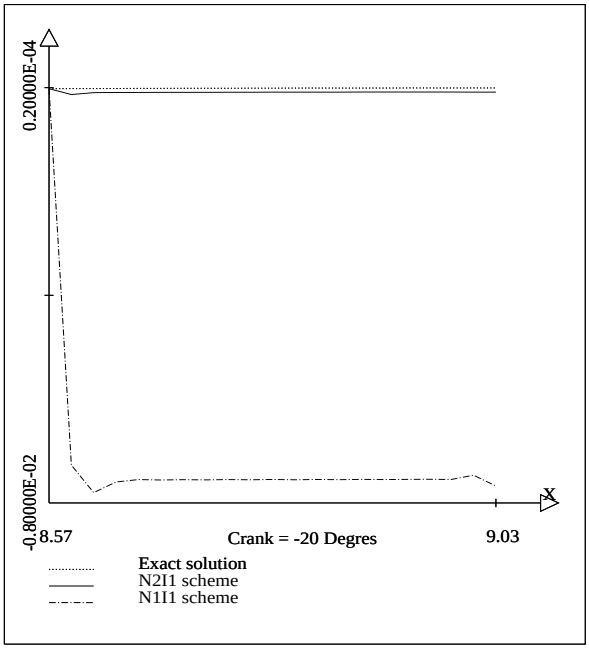


Figure 9: Isentropic compression : density

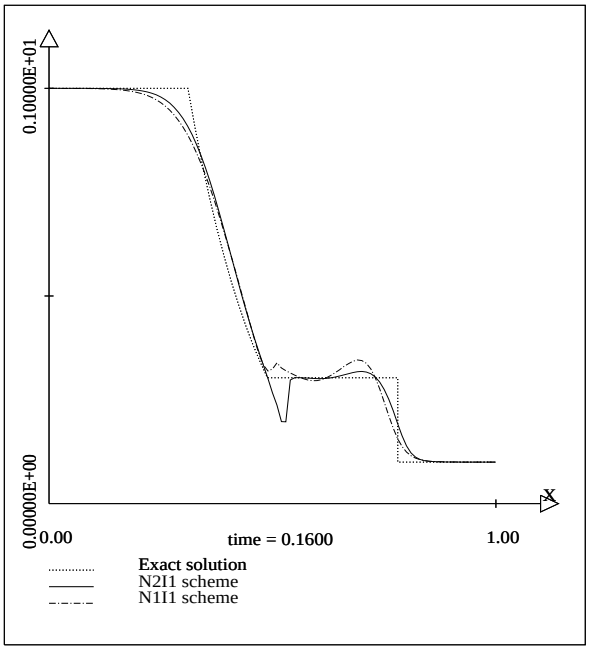


Figure 10: Sod Shock Tube : pressure

Albada van Leer limiters have been used, but we do not employ any "entropy fix". Figures (10,11,12) display respectively the pressure, density and velocity at time $t = 0.16$. The most remarkable feature of the N2I1 scheme is the large oscillation present at the rear of the rarefaction wave. This oscillation is much more smaller with N1I1 scheme. On this other hand, the N2I1 scheme more faithfully respects the contact discontinuity (11) which is barely visible on the results obtained with the N1I1 scheme.

5 Flow in a bi-dimensional model of a piston engine.

We end this paper by the study of the flow in the intake phase of a bi-dimensional planar model of a piston engine. The geometry under investigation

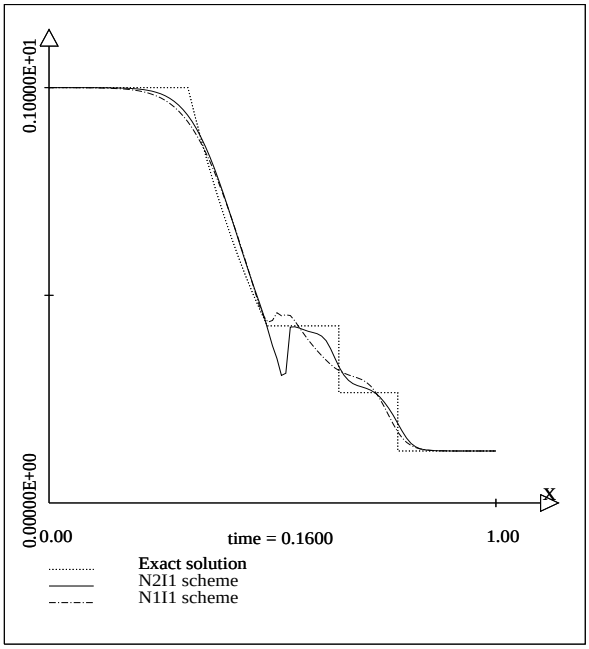


Figure 11: Sod Shock Tube : density

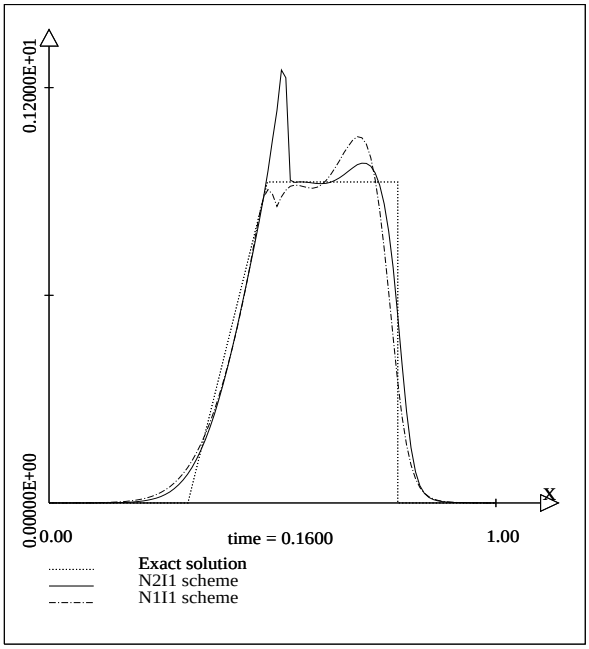


Figure 12: Sod Shock Tube : Velocity

is given Figure 13.a and 13.b at two different times in the cycle. The diameter of the cylinder is 8.8 cm, its heigh is 9.032 cm at bottom dead center (BDC) while the squish region at top dead center (TDC) is 0.132 cm high. The diameter of the intake pipe is 3.06 cm. In this computation, the concurrent motion of the piston and of the valve have to be considered. Figure 14 displays the displacement law of the piston and valve with respect to the crank angle (CA). The computation extends from the opening of the valve at -10^oCA to its closing at 230^oCA. Figure 14 shows that the piston is still moving upward at the beginning of the valve opening. The valve begins its closure movement at 110^oCA and is totally closed at 230^oCA while the piston begins to move upward at 180^oCA. In the present computation, the valve is considered closed when it is 1 mm close from the cylinder head. At the beginning of the computation, we consider that the flow is at rest with pressure and density equal to their atmospheric value i.e $P_{atm} = 10^6 cgs$ and $\rho_{atm} = 1.191^{-3} g.cm^{-3}$. We use slip condition at the boundaries except at the intake pipe where the value of the flux is given through a Steger-Warming formulation :

$$\Phi_{i\Gamma} = A^+(W_i)W_i + A^-(W_i)W_{atm}$$

where W_{atm} is the atmospheric state at rest : $W_{atm} = (\rho_{atm}, 0, 0, P_{atm})^t$, A^+ (resp A^-) is the positive (resp negative) part of

$$A = \partial_W(n_x F(W) + n_y G(W))$$

where

$$\vec{n} = (n_x, n_y)$$

is the outward unit normal.

5.1 Mesh movement.

The concurrent motions of the valve and piston impose to move the mesh in an adequate way. Because these two motions do not have the same translation axis, it is difficult to specify the movement of the mesh in an analytical way. So instead we compute the motion of the mesh by solving the following laplacian

problem : Let (u,v) be the speed of each mesh points. Then we solve :

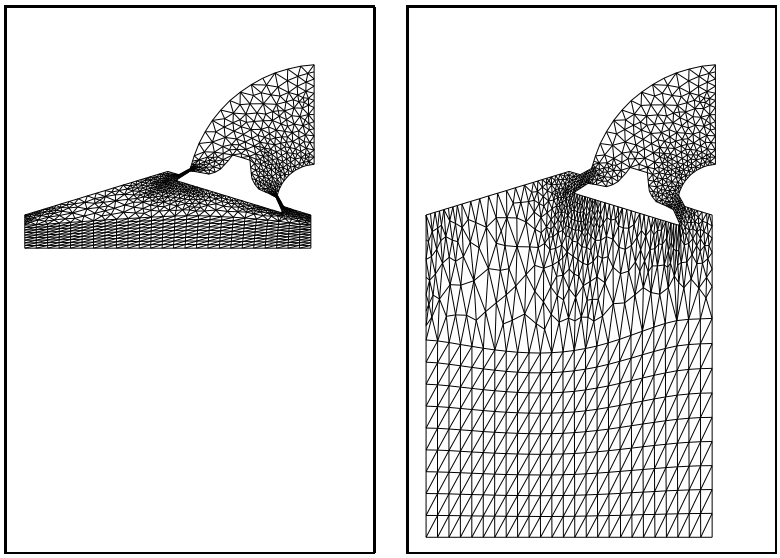
$$\left\{\begin{array}{lcl|lcl} \Delta.u & = & 0 \text{ on } \Omega & \Delta.v & = & 0 \text{ on } \Omega \\ & & & \frac{\partial v}{\partial n} & = & 0 \quad \text{on the cylinder sides} \\ u & = & f \text{ on } \partial\Omega & v & = & g \text{ on } \partial\Omega/\text{the cylinder sides} \end{array}\right. \tag{31}$$

Although not totally satisfactory (in particular, this algorithm does not exclude possible crossing of mesh points), this technique has been found sufficient for this case. Problem (31) is solved in an efficient way by a conjugate gradient algorithm and represents only a very modest part of the total computational time.

5.2 Results

The purpose of this section is to compare the results obtained with the second order DeC scheme (N2I1) with those obtained with the first-order implicit scheme (3) obtained using a first-order time accurate discretization formula. Thus two computations have been performed on the same mesh displayed in Figure 13 with two different time steps. Figure 15.a shows the mass flow versus the crank angle obtained with the two schemes. In these computations, the time step is fixed and corresponds to a variation of 5^o of the crank angle. It can be noticed that there is almost no difference between the results obtained by the two schemes. Figure 15.b displays the results obtained with a time step corresponding to a variation of 10^oCA. Again there is almost no difference between these two curves and the previous two ones. This indicate that the time and space accuracies are not critical for the computation of integrated quantities like the mass flow. However if one is interested in the detail of the motion of the flow, large differences can now be noticed in the results obtained by the two schemes. For instance, Figure 16 displays the streamlines at 230^oCA obtained with a time step of 5^oCA. The organization of the motion is on the average the same for the two schemes : Two large recirculating regions are present under the two sides of the valve. But the centers of the vortices are higher in the results obtained with the second-order scheme and the vortex on the right side of the intake valve is much more developped. These differences are emphazised if one looks in the flow around the intake valve (Figure 17.a and 17.b). It can be seen that the second order scheme (N2I1) predicts the

existence of two recirculating regions on the upper side of the intake valve while these region are barely visible in the results obtained with the first order scheme. Note also that the position of the line of separation between the two recirculating region under the valve is very different : It is almost symmetrical with respect to the valve in Figure 17.b while it is shifted on the left in Figure 17.a.
Figure 18 now shows the results obtained by the second-order scheme (N2I1) with a time step of 10^oCA. It is interesting to notice that these results are closer to the results obtained with (N2I1) and a time step of 5^oCA than are the results obtained with the first order scheme and a time step of 5^oCA. In particular, the existence of two recirculating regions on the upper part of the intake valve is again predicted and the position of the line of separation is again shifted on the left. This indicate the importance of the spatial accuracy for this type of computation. It is worth mentionning that it has not been possible to use a time step of 10^oCA with the first order scheme. The computation exploded right before 230^o CA (this explains the vertical part of the curve displayed in Figure 15.d). It is likely that the second Newton iteration performed in the (N2I1) scheme reinforces the stability of the scheme. Therefore at least for this case, the use of a second-order scheme not only improves the accuracy of the results but also has a beneficial influence on the stability of the computation.



a) Position at 0^oCA b) Position at 180^oCA

Figure 13: Time evolution of the unstructured mesh

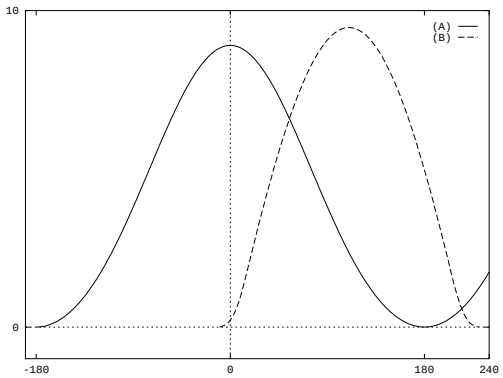


Figure 14: Position versus Crank angle : (A) Piston motion in cm - (B) Intake Valve motion in mm

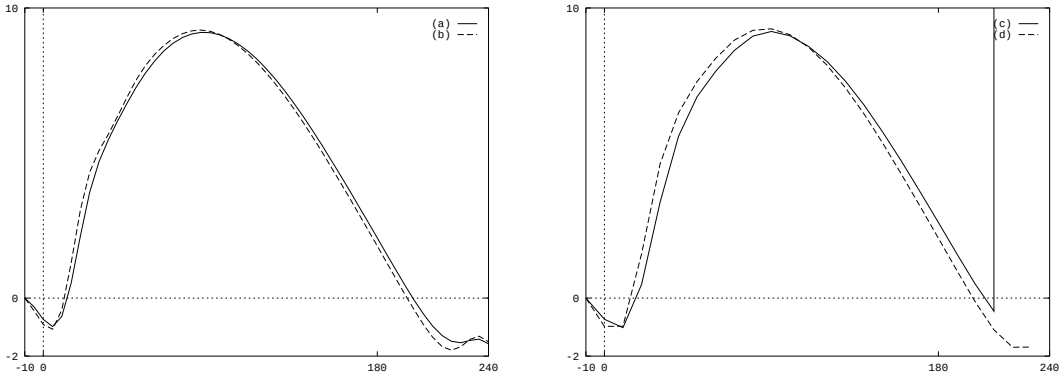
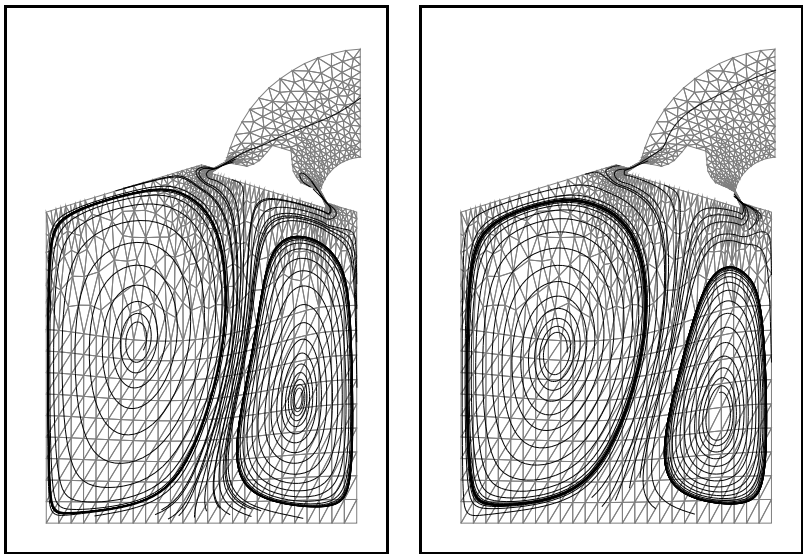
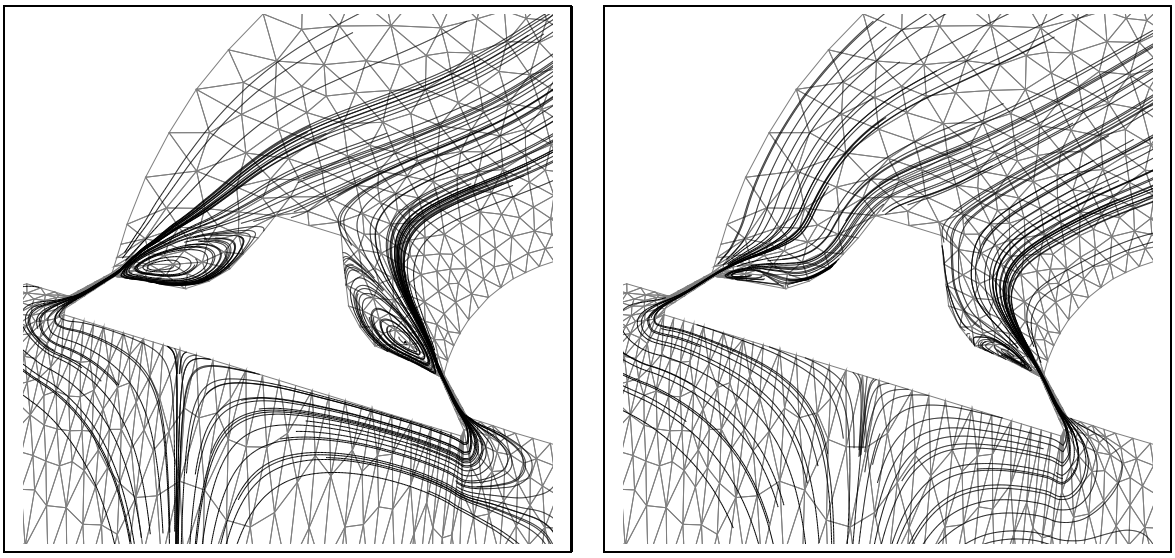


Figure 15: Mass flow rate versus Crank angle : (a) First order scheme with $\delta\theta = 5^oCA$ - (b) N2I1 scheme with $\delta\theta = 5^oCA$ - (c) First order scheme with $\delta\theta = 10^oCA$ - (d) N2I1 scheme with $\delta\theta = 10^oCA$



a) N2I1 scheme b) First order scheme

Figure 16: Stream Lines at 230^oCA with $\delta\theta = 5^oCA$



a) N2I1 scheme b) First orderS1/2T1 scheme

Figure 17: Stream lines near the intake valve at 230^oCA with $\delta\theta = 5^oCA$

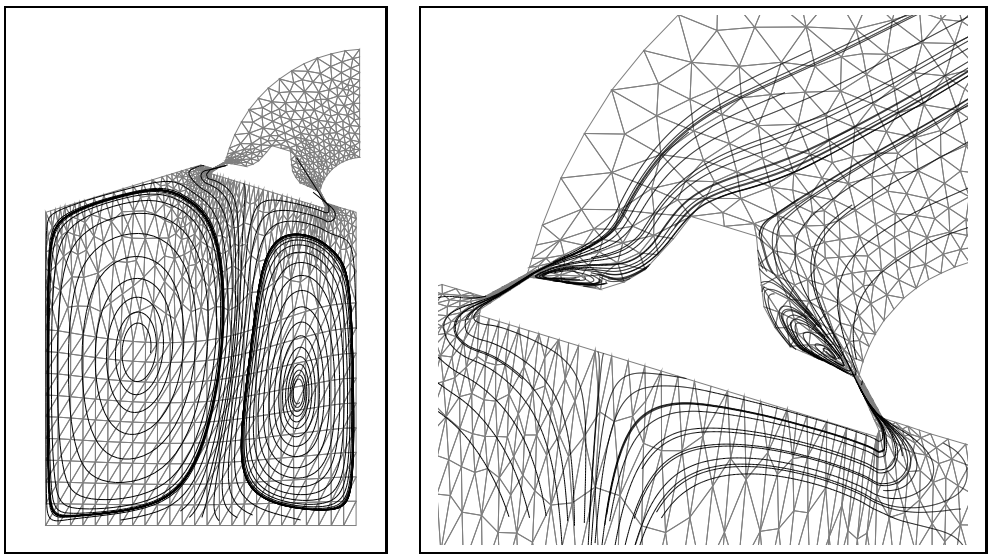


Figure 18: N2I1 scheme - Stream Lines at $230^o CA$ with $\delta\theta = 10^o CA$

6 Conclusion

This work proposes a general method to construct high order implicit schemes at the price of inverting only a finite number of linear systems coming from first order schemes. For the case of second order, we demonstrated that this number can be as small as one, while preserving the linear stability of the scheme. In practice, however, we have found that in some case, it is better to perform two iterations of the scheme (12), and we recommend in general this option (Although in some cases, scheme N1I1 gives better results than N2I1 see section 4.4). We emphasize that when an iterative method is used to solve the linear systems, this does not imply that the computational time is doubled: due to the good initialisation of the iterative solver provided by the first step, the second linear system is much more simpler to solve than the first one. In practice, we experimentally found that scheme N2I1 is only less than 50% more expensive than scheme N1I1. The numerical experiments reported in this paper conclusively show that these Dec schemes provide a very interesting way to construct high order implicit method.

Acknowledgment

Several points in this work have been discussed with J.A. Désidéri. We also thank P.Hemker who attracts our attention on the different ways to initialize the Dec iterations in connection with backward differentiation formula.

This study has been supported in the context of a cooperation between Renault S.A., Simulog and INRIA in order to develop efficient numerical methods for the simulations of the aerodynamicsof combustion chambers.

References

[1] P ARMINJON and A DERVIEUX. Construction of tvd-like artificial viscosities on two-dimensional arbitrary fem grids. *J. Comp. Phys.*, 106:176–198, 1993.

[2] R.M. BEAM and R.M. WARMING. An implicit finite difference algorithm for hyperbolic system in conservation law form. *J. Comp. Phys.*, 22:67–110, 1976.

[3] J.A. DESIDERI and P.W. HEMKER. Analysis of the convergence of iterative implicit and defect-correction algorithms for hyperbolic problems. *INRIA report*, 1200, 1990.

[4] L. FEZOUI. Résolution des équations d'euler par un schéma de van leer en éléments finis. *INRIA report*, 358, 1985.

[5] L FEZOUI and B STOUFFLET. A class of implicit upwind schemes for euler simulations with unstructured meshes. *J. Comp. Phys.*, 84:174–206, 1989.

[6] W. HACKBUSH. *Multi-Grid Methods and Applications.* Springer Verlag, 1985.

[7] A. LERAT, J. SIDES, and V. DARU. An implicit finite volume method for solving the euler equations. *Lecture Notes in Physics*, 170:343–349, 1982.

[8] B. NKONGA. *Développement de méthodes numériques pour les écoulements tridimensionnels réactifs dans un domaine déformable.* Thèse de l'université de NiceSophia-Antipolis, 1992.

[9] B. NKONGA, G. FERNANDEZ, H. GUILLARD, and B. LARROUTUROU. Numerical investigations of the tulipe flame instability - comparisons with experimental results. *Combust. Sci. and Tech.*, 87:69–89, 1992.

[10] B. NKONGA and H. GUILLARD. Godunov type method on non-structured meshes for three-dimensional moving boundary problems. *Comput. Methods Appl. Mech. Engrg.*, 113:183–204, April 1993.

[11] R. PEYRET and T.D. TAYLOR. *Computational Methods for Fluid Flow.* Springer-Verlag, 1983.

[12] R.D. RICHTMYER and K.W. MORTON. *Difference methods for initial-value problems.* Interscience, 1967.

[13] Ph. ROE. Approximate riemann solvers, parameter vectors, and difference schemes. *J. Comp. Phys.*, 43:357–371, 1981.

[14] G.D. VAN ALBADA, B. VAN LEER, and W.W. ROBERTS. A comparative study of computational methos in cosmic gas dynamics. *Astronomy ans Astrophysics*, 108:76–84, 1982.

[15] B. VAN LEER. Towards the ultimate conservation difference scheme v, a second-order sequel to godunov's method. *J. Comp. Phys.*, 32:101–136, 1979.

[16] H.C. YEE. Upwind and symmetric shock capturing schemes. *NASA Tech. Memorandum 89464*, May 1987.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399